Exam Paper: DES-PSP Presidential Concept Stock Price Prediction via Double Encoder Seq2Seq Structure

Final Exam of Computational Intelligence

Zijin Hong
2020101911
Mathematics and Applied Mathematics
Jinan University - JBJI
hongzijin@stu2020.jnu.edu.cn

1 Abstract

Stock price prediction is of paramount importance for financial decision-making. Even with strides made through deep learning, the task remains challenging due to the complexity of historical information and the impact of legal policies and societal factors. Presidential concept stocks represent the stock sectors expected to benefit from presidential elections. The president exerts substantial influence over national policies and the economy - changes that directly affect the stock performance of relevant businesses. By analyzing the influence of concept stocks during past elections, we can predict the impact of future elections on stock prices. The key to concept stock prediction is capturing the price fluctuations during election results. In this paper, we introduce a novel deep learning model: **DES-PSP**, based on the structure of Seq2Seq, consisting of Double Encoders and a Fluctuation Decoder. The Double Encoders, including the President Encoder and Competitor Encoder, capture the feature of stock price based on the study of various concept stocks of presidential candidates. The Fluctuation Decoder aims to utilize the stock price movement to guide the prediction. Compared to other RNN-based models, DES-PSP makes accurate predictions based on policies and societal factors while sensitively capturing series trends. In order to verify the efficiency and accuracy of DES-PSP, we conducted comparative experiments with several advanced baseline models. The experimental results demonstrate that DES-PSP is superior to them effectively.

2 Introduction

Time series prediction is a fundamental task in many fields, aiming to predict future values based on historical data. The approach for such predictions typically involves identifying meaningful patterns from past observations and leveraging these patterns to predict future trends. With wide-ranging applications in various domains like finance, economics, weather forecasting, and demand forecasting, this technique has gained significant attention.

In earlier research, simple single-series linear models such as ARIMA were predominantly used for time series prediction [1]. While effective in simpler scenarios, these models often struggle with the nonlinear relationships and overlooked external influencing factors in time series data, limiting their predictive accuracy in complex environments [2, 3, 4].

With the advent of deep learning, more sophisticated methods have emerged that can handle complex and nonlinear relationships, as well as the underlying interdependencies present in multivariate time series. Neural network models such as Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU) have demonstrated impressive performance in time series prediction tasks [5, 6]. However, despite their ability to manage long-term dependencies, these models often struggle with dynamic changes and noises at different time scales [7, 8].

Attempts to address these challenges have seen the exploration of other types of models, such as Convolutional Neural Networks (CNNs) and Transformers [9, 10, 11, 12, 13, 14]. These models can encode more complex temporal patterns and capture long-term dependencies within larger context windows. However, they often neglect the impact of external events, crucial in domains like stock market prediction, where political, economic, and social events can significantly affect the trend.

To overcome these challenges, we propose a **D**ouble **E**ncoder **S**eq2seq-Based **P**residential Concept **S**tock Price **P**rediction model (**DES-PSP**), specifically designed for stock market prediction during the U.S. presidential elections. By categorizing companies into "Trump stocks" and "Biden stocks" based on their associated favorable policies, we aim to capture the impact of political events on stock market dynamics. Our model features a novel dual-encoder structure that separately encodes the stock price data of the identified presidential concept stocks and the rest of the market. This structure allows for more accurate predictions by specifically dealing with the influence of political events on these stocks. Additionally, our model introduces the Fluctuation Decoder to handle complex historical information better, capturing changes around critical turning points often overlooked in traditional time series prediction models. The main contributions of this paper are summarized as follows:

- We propose a novel seq2seq-based deep learning model specifically designed for stock price prediction during U.S. presidential elections. By categorizing companies based on the policies of specific presidential candidates, we attempt to unveil the effects of presidential elections on the stock market dynamics.
- Our model introduces a dual-encoder structure that independently encodes the stock price data of the identified concept stocks and the rest of the market. This structure allows us to more accurately predict price movements after the election, dependent on the winning candidate.
- We innovate a Fluctuation Decoder mechanism to better handle complex historical information and capture changes around critical turning points, contributing to improved prediction accuracy.

3 Proposed Method

In this section, we will introduce our proposed method in three aspects. First, we will discuss the relative definitions and statements of DES-PSP. Secondly, we will introduce the data preprocessing process in our model. Finally, we will comprehensively describe our proposed model, DES-PSP, including the Double Encoder and Fluctuation Decoder. Note that since people start paying attention to the presidential election very early and use the predicted stocks to make potential ultra-short, short and mid-range trades, we focus on predicting the sequence of two week's prices by using a sequence of the previous year in this paper.

3.1 Definitions and Problem Statements

DES-PSP aims to predict the concept stock price of the current elected president of the American election based on the information from the last election. To more clearly express the relevant operations in our method, the definitions we need to introduce are as below:

Definition 1.*(Concept Stock Prediction):* The election of different presidents will lead them to bring different favorable policies. Presidential concept stocks refer to the U.S. stock sectors expected to benefit from the election of each presidential candidate. In order to predict the current presidential stock, we need to analyze the former presidential stock of the last election. And then the key to Concept Stock Prediction is to capture the variation of stock price during the elections, which means studying the former election to predict the current one.

Definition 2.*(Presidential Concept Stock & Competitor Concept Stock):* The current presidential concept stock refers to the stocks favored by the elected presidential candidate that we would like to predict. In addition, the former presidential concept stocks refers to the stocks favored by the president who won the previous presidential election. The competitor concept stocks, by contrast, are those of presidential candidates who lost in the previous election. We assume that the stocks we want to predict correspond to the presidential victory and analysis made combined with the previous election.

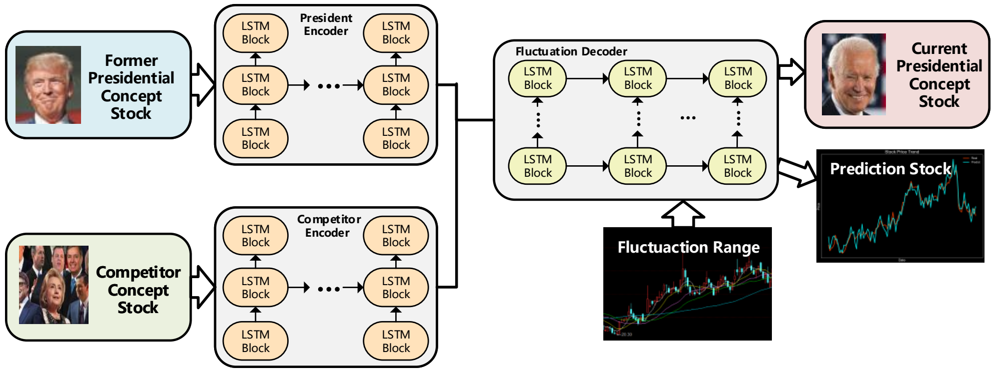

Figure 1: The framework of DES-PSP

Definition 3.*(Trump Concept Stock & Biden Concept stock):* In this paper, we aim to predict the presidential concept stock of Joe Biden during 2020 by studying the concept stocks of the election 2016, where we set Donald Trump concept stock as former presidential concept stock, and the other competitor of him we set it to competitor concept stock. The following several investment bank analysts analyzed Biden and Trump winning the election will be positive for U.S. stock sectors:

- Biden Concept Stocks: Alternative energy stocks, Infrastructure stocks, Small cap stocks, Export leading companies, Cannabis stocks.
- Trump Concept Stocks: Banking stocks, Oil stocks, Technology leader's stocks, Defense stocks, Pharmaceutical stocks.

3.2 Data Preprocessing

We have obtained raw stock price data from the entire U.S. stock market spanning January 2^{nd}, 1995 to February 24^{th}, 2023, including opening, highest, lowest, closing prices, volume, fluctuation amount, and fluctuation range (%). The pre-vesting closing price serves as the feature in our prediction task. Unlike conventional single stock price trends, DES-PSP, a Seq2Seq model, is sensitive to changes around elections and aims to predict concept stock levels. The data preparation process entails four stages. First, stock price data is divided into four categories: Biden Concept Stocks, Trump Concept Stocks, Trump's Competitor's Concept Stocks, and others. Second, invalid data, like delisted or unlisted stocks within our target window, are removed. Then, the data is divided into training and prediction periods based on election dates. Finally, the price data is standardized using the Standard Normalization method.

3.3 DES-PSP

The ubiquitous Seq2Seq model consists of an encoder (Enc) transforming the input sequence (X) into hidden states (h), and a decoder (Dec) mapping these states to the output (Y). Our proposed concept stock price prediction model, DES-PSP, extends this Seq2Seq structure, introducing Double Encoders and a Fluctuation Decoder. The Double Encoders, including a President Encoder and a Competitor Encoder, aim to capture comprehensive stock information via different encoders, extracting multiple change features from both favorable and regular stocks. The Fluctuation Decoder uses each stock's fluctuation range to guide predictions, thus enhancing the model's predictive capabilities.

3.3.1 Double Encoder

In time-series problems, Recurrent neural networks are widely utilized in time-series problems, especially the variant LSTM, which was designed with a forgetting gate and a memory cell; the network ensures that LSTM can learn the long-term temporal dependencies on those time-series data. An LSTM network generally combines several parts, including input, hidden state, cell memory, model layer, Etc. The detailed process of LSTM can be represented by the formula below:

$$f_t = Sigmoid(W_f h_{t-1} + W_f X_t + b_f) \tag{1}$$

$$i_t = Sigmoid(W_i h_{t-1} + W_i X_t + b_i) \tag{2}$$

$$\tilde{C}t = tanh(W_c h_{t-1} + W_c X_t + b_c) \tag{3}$$

$$O_t = Sigmoid(W_o h_{t-1} + W_o X_t + b_o) \tag{4}$$

$$C_t = f_t C_{t-1} + i_t \tilde{C}_t \tag{5}$$

$$h_t = O_t tanh(C_t) \tag{6}$$

where f_t, i_t, C_t, and h_t denote the forget gate, input gate, cell memory, and hidden state at the time of t, and x_t, C_{t-1}, and h_{t-1} are the input data at the time of $t-1$; two activate functions tanh and sigmoid are given, while W and b are the learnable parameters in LSTM.

While favorable presidential policies do not guarantee an uninterrupted uptrend in associated concept stocks, the Double Encoders in DES-PSP are designed to encapsulate more information from various stocks, curbing overfitting. Each encoder embeds multiple features from both favorable and regular stocks concurrently, thus enhancing feature representations. The five-layer LSTM network is used for both encoders. The President Encoder focuses on the stock price features of the previous election winner, while the Competitor Stock Encoder processes the price information of less-influential or unfavorable stocks. This dual-encoder design allows for a more balanced and comprehensive forecast.

The President Encoder E_{pre}, which is specifically designed to embed the former election winner stock price features, inputs the data consisting of normalized former presidential concept stock $X_{pre} \in \mathbb{R}^{365\times 1}$ into the network and output outputs the hidden state h_{pre} and cell memory c_{pre}:

$$h_{pre}, c_{pre} = E_{pre}\left(X_{pre}\right) \tag{7}$$

The Competitor Stock Encoder, which deals with the information for the other price, such as unfavorable stock and the stock without significant influence, uses the competitor concept stock $X_{com} \in \mathbb{R}^{365\times 1}$ as the input to the network, similarly outputting hidden state h_{com} and cell memory c_{com}:

$$h_{com}, c_{com} = E_{com}\left(X_{com}\right) \tag{8}$$

Note that h_{pre}, c_{pre}, h_{com}, and $c_{com} \in \mathbb{R}^{layers\times batchsize\times hidden}$. Finally, for a better fusion of features, we add the hidden states and the cell memories of the two encoders with a competitor ratio α to obtain the final result:

$$h_{enc} = h_{pre} + \alpha h_{com} \tag{9}$$

$$c_{enc} = c_{pre} + \alpha c_{com} \tag{10}$$

3.3.2 Fluctuation Decoder

The fluctuation range (FR) is the percentage of a stock's price change the next day divided by the previous day's price. The positive value of the fluctuation range represents the price rise, and the negative refers to the fall. The Fluctuation Decoder is targeted to utilize the range of fluctuation of each stock to guide the prediction. In this way, we can better express the correlation between the movements of stocks and the price changes. The Fluctuation Decoder is composed of five LSTM layers. The decoder takes the Fluctuation range (FR), presidential concept stock price, h_{enc}, and c_{enc}, generated by the Double Encoders as inputs. Finally, the model outputs the predicted price of 14 days.

The Seq2Seq-based model can predict multiple points instead of using a sliding window mechanism. We use y_{t-1} as the previous stock price, when predicting y_t at the time of t during training:

$$M_t = y_{t-1} + FR \tag{11}$$

$$\widehat{M_t} = tanh(W M_t + b) \tag{12}$$

$$\widehat{y}_t, (h_{dec}, c_{dec}) = LSTM(\widehat{M}_t, h_{enc}, c_{enc}) \tag{13}$$

where $LSTM$ denotes the LSTM layer, W, b refer to the parameters of full connected layer. Which can be reduced to:

$$\widehat{y}_{t-1} = dec(y_{t-1}, FR, h_{enc}, c_{enc}) \tag{14}$$

where dec denote the mapping function of the Fluctuation Decoder.

4 Experiments

4.1 Datasets

We conducted experiments based on a dataset that we created and annotated ourselves. The dataset includes 8514 stocks of U.S. stocks from January 2^{nd},1995 to February 24^{th}, 2023, and high-quality four classifications of stocks based on the hypothetical scenario of Biden or Trump winning the election. We used Trump concept stocks as the Former Presidential Concept Stocks training set and the Current also Biden concept stocks as the test set. The former includes 3119 stocks, and based on the closing price of these stocks from November 7^{th}, 2015 to November 7^{th}, 2016, we predicted the stock prices of these stocks within 14 days after Trump won the election. The latter includes 3439 stocks, which were used to verify the accuracy of the model's predictions of stock prices after Biden won the election. The competitor concept stocks corresponding to Trump concept stocks are in the same period. The number is 4124; we randomly remove some to balance them with Trump concept stocks.

4.2 Baselines

To evaluate the performance of DES-PSP, we juxtapose it against various benchmark models, including both traditional and deep learning time-series models. The traditional models encompass ARIMA[15], a statistical forecasting model, and VAR[16], a model suited for multivariate time-series prediction. In addition to the methods above, RNN baseline models are utilized in a sliding window mechanism, such as 5-layer LSTM, 5-layer BiLSTM, and similarly GRU, and BiGRU are also utilized. The evaluation also includes Seq2Seq structure models such as LSTM–LSTM, BiLSTM–LSTM, GRU–GRU, and BiGRU–GRU. The fundamental difference between the baseline LSTM–LSTM and DES-PSP is the input - while the former employs a single grouped stock price, the latter includes multiple groups features.

4.3 Experimental Results

Three metrics are used to evaluate the performance of the model in our experiments: Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and Final Displacement Error (FDE). In our experiments, we use Adam as the optimizer and set the learning rate to 0.001, the competitor ratio to 0.3, the mini-batch size to 512, the hidden size to 64, and the number of epochs to 100. Before the input, we normalized the data by standard normalization. Table 1 shows the comparison results of different methods.

Table 1: Comparison results with baselines

Model	**RMSE**	**MAE**	**FDE**
ARIMA	1.9733	1.124	1.274
VAR	3.6431	2.358	3.702
GRU	0.7663	0.5512	0.0712
BiGRU	0.7652	0.5506	0.0704
LSTM	0.7661	0.5544	0.0697
BiLSTM	0.7591	0.5488	0.0607
GRU–GRU	0.8061	0.5994	0.0883
BiGRU–GRU	0.7993	0.6089	0.0861
LSTM-LSTM	0.7542	0.5439	0.0554
BiLSTM–LSTM	0.7337	0.4979	0.0495
DES-PSP	**0.6559**	**0.4190**	**0.0373**

As can be seen from the table, our proposed model is superior to other models in the three evaluation metrics involved. We attribute these results to our proposed model's structure take into account a much richer set of features of social factors, and the involvement of fluctuation range makes our model better able to capture complex historical information. Unlike time-series models utilizing stock price only, DES-PSP benefits from more information, including the features of the regular stocks and the unfavorable stocks of the incoming policies and the guided fluctuation range. The former factor improves the model's generalization when predicting the stock price of a large base, and the latter enhances the ability to predict price trends; that is, even if there is no precise price prediction, it can accurately judge the rise or fall of the stock price.

4.4 Further Analysis

In order to verify the impact of the president's favorable policies on the stock price considered in the presidential concept stock price prediction, we canceled the concept stock classification, used the total amount of market stock data of the same period of Trump concept stocks as the training data, and still made the prediction on the Biden concept stocks. We use RMSE as the evaluation metrics with the same parameters above the experiment. The result is shown below:

Table 2: Comparison of different prediction task under different metrics

Model	Concept Stocks	Total amount Stocks
LSTM	0.7661	0.9817
BiLSTM	0.7591	0.9660
LSTM-LSTM	0.7542	0.9558
BiLSTM-LSTM	0.7337	0.9457

From Table 2, we can see the concept stock price prediction has a significant positive impact on stock price prediction, the classification of concept stocks filters out some useless information for predicting target stocks, and taking into account the impact of social factors and policies when making predictions that effectively improves the accuracy of the whole prediction task.

5 Conclusions

As the remaining challenges of stock price prediction, considering the features of the complexity of historical information and the impact of legal policies and societal factors become the way to improve accuracy. In response to the above challenges, we propose a novel stock price prediction DEP-PSP consisting of Double Encoders and a Fluctuation Decoder in this paper. To verify the improvement of the model, we conducted comparative experiments on several baseline models; the results show that the proposed DEP-PSP model outperforms others effectively. However, it still has some limitations. The classification of concept stocks requires much manual labeling. In addition, DES-PSP needs further verification when dealing with more complex scenarios and other concepts apart from presidential ones.

References

[1] George E.P. Box, Gwilym M. Jenkins, and G. Reinsel. Time series analysis: Forecasting and control Holden-day San Francisco. *BoxTime Series Analysis: Forecasting and Control Holden Day1970*, pages 199–201, 1970.

[2] Tim Bollerslev. Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 31:307–327, 1986.

[3] Rob J. Hyndman and Yeasmin Khandakar. Automatic time series forecasting: The forecast package for R. *Journal of statistical software*, 27:1–22, 2008.

[4] George EP Box, Gwilym M Jenkins, Gregory C Reinsel, and Greta M Ljung. *Time series analysis: forecasting and control*. John Wiley & Sons.

[5] Jia Wang, Tong Sun, Benyuan Liu, Yu Cao, and Hongwei Zhu. CLVSA: A convolutional LSTM based variational sequence-to-sequence model with attention for predicting trends of financial markets. *arXiv preprint arXiv:2104.04041*, 2021.

[6] Subba Rao Polamuri, Kudipudi Srinivas, and A. Krishna Mohan. Multi-model generative adversarial network hybrid prediction algorithm (MMGAN-HPA) for stock market prices

prediction. *Journal of King Saud University-Computer and Information Sciences*, 34:7433–7444, 2022.

[7] Zexuan Yin and Paolo Barucca. Stochastic Recurrent Neural Network for Multistep Time Series Forecasting. In Teddy Mantoro, Minho Lee, Media Anugerah Ayu, Kok Wai Wong, and Achmad Nizar Hidayanto, editors, *Neural Information Processing*, Lecture Notes in Computer Science, pages 14–26, 2021.

[8] Yuntao Du, Jindong Wang, Wenjie Feng, Sinno Pan, Tao Qin, Renjun Xu, and Chongjun Wang. AdaRNN: Adaptive Learning and Forecasting of Time Series. In *Proceedings of the 30th ACM International Conference on Information & Knowledge Management*, CIKM '21, pages 402–411, New York, NY, USA, 2021.

[9] Jimmy Ming-Tai Wu, Zhongcui Li, Gautam Srivastava, Jaroslav Frnda, Vicente Garcia Diaz, and Jerry Chun-Wei Lin. A CNN-based stock price trend prediction with futures and historical price. In *2020 International Conference on Pervasive Artificial Intelligence (ICPAI)*, pages 134–139, 2020.

[10] Jimmy Ming-Tai Wu, Zhongcui Li, Norbert Herencsar, Bay Vo, and Jerry Chun-Wei Lin. A graph-based CNN-LSTM stock price prediction algorithm with leading indicators. *Multimedia Systems*, pages 1–20, 2021.

[11] S. Kumar Chandar. Convolutional neural network for stock trading using technical indicators. *Automated Software Engineering*, 29:1–14, 2022.

[12] Zekai Chen, Dingshuo Chen, Xiao Zhang, Zixuan Yuan, and Xiuzhen Cheng. Learning graph structures with transformer for multivariate time-series anomaly detection in IoT. *IEEE Internet of Things Journal*, 9:9179–9189, 2021.

[13] Amin Shabani, Amir Abdi, Lili Meng, and Tristan Sylvain. Scaleformer: Iterative multi-scale refining transformers for time series forecasting. *arXiv preprint arXiv:2206.04038*, 2022.

[14] Ranak Roy Chowdhury, Xiyuan Zhang, Jingbo Shang, Rajesh K. Gupta, and Dezhi Hong. TARNet: Task-Aware Reconstruction for Time-Series Transformer. In *Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining , Washington, DC, USA*, pages 14–18, 2022.

[15] Shaun S Wulff. Time series analysis: Forecasting and control. *Journal of Quality Technology*, 49(4):418, 2017.

[16] John C Robertson and Ellis W Tallman. Vector autoregressions: forecasting and reality. *Economic Review-Federal Reserve Bank of Atlanta*, 84(1):4, 1999.